Autocorrelation and Regularization in Digital Images II. Simple Image Models

DAVID L. B. JUPP, ALAN H. STRAHLER, MEMBER, IEEE, AND CURTIS E. WOODCOCK

Abstract—**The variogram function used in geostatistical analysis is a useful statistic in the analysis of remotely sensed images. In Part II of this paper, using the results derived in Part I, the basic second-order, or covariance, properties of scenes modeled by simple disks of varying size and spacing after imaging into disk-shaped pixels are analyzed to explore the relationship between image variograms and discrete object scene structure. The models provide insight into the nature of real images of the earth's surface and the tools for a complete analysis of the more complex case of three-dimensional illuminated discrete-object images.**

Keywords—**Remote sensing, digital image, spatial statistics, autocovariance, regionalized variable, variogram.**

I. INTRODUCTION

IN MANY applications of remote sensing and digital image processing, it is reasonable to regard a scene as a spatial arrangement of two- or three-dimensional objects superimposed upon a background [1]. Consider a set of measurements of brightness values obtained from the scene, each of which is associated with a spatial position in the scene. In mathematical terms, this image of the scene can be considered as a "regionalized variable"—brightness as a function of spatial position. In the first paper of this series [2], we considered the covariance properties and effects of regularization on scenes modeled by general regionalized variables. We also introduced a range of tools allowing us to construct covariance functions for discrete objects. The objective of this paper is to apply the results of the first paper to some simple models of scenes and digital images obtained from them, in order to show how the spatial structure of an image depends on the underlying spatial structure of the scene. We begin with an introduction to the regionalized variable and the variogram.

Consider a spatial random function f (regionalized variable) that generates radiance values $f(x)$ as a function of spatial position x. (Table I presents a description of the symbols used in this paper.) Here x is regarded as a vector variable which, in our case, is a pair of position coordinates. In geostatistics, the degree of spatial autocovariance associated with a regionalized variable is commonly analyzed using the properties of the variogram function [3]–[7]

$$V(h) = \tfrac{1}{2}E\left\{\left(f(x) - f(x+h)\right)^2\right\} \qquad (1)$$

where $V(h)$ is the value of the variogram function at (vector) distance h, $f(x)$ is the value of the regionalized variable at (vector) location x, and $f(x+h)$ is the value of the regionalized variable at location $x+h$. The variogram and the spatial covariance are simply related by

$$V(h) = \mathrm{Cov}\,(0) - \mathrm{Cov}\,(h) \qquad (2)$$

TABLE I
NOTATION

Symbol	Description
$f(x)$	Spatial random function
$V(h)$	Variogram at distance h
$\mathrm{Cov}(h)$	Covariance of $f(x)$ at h
$a_n(V)$	Serra's range quantity
$\mathbf{R}^N$	Parameter space of N dimensions
$\mathbf{Z}$	Regularization element
$f_{\mathbf{Z}}(y)$	Spatial random function regularized by $\mathbf{Z}$
$\mathrm{Mes}(\mathbf{Z})$	Lebesque measure of $\mathbf{Z}$; area of $\mathbf{Z}$ for our case
$I_{\mathbf{Z}}$	Indicator function for $\mathbf{Z}$
$V_{\mathbf{Z}}(h)$	Regularized variogram
T	Normalized overlap function for $\mathbf{Z}$
Q_1	Disk model proportion of uncovered area
Q_2	Disk model set overlapping function for uncovered area
$\mathbf{D}$	Disk model area covered by disks
$K(h)$	Set overlap function for individual disks
$\gamma(h)$	Disk model variogram
$\gamma_{\mathbf{Z}}$	Regularized disk model variogram
$\sigma_{\mathbf{Z}}^2$	Far-field variance for disk model
D_1, D_2	Diameters of basic and regularizing disks
$C_{\mathbf{Z}}$	Far-field regularized disk model variance
$G(u)$	Generating function for the point process
P	Crown fraction

Manuscript received August 4, 1987; revised November 1, 1988. This work was supported by NASA under Grants NAG 5-273, NAG 5-276, NAS 9-16664 (Subcontract L200080), NAGW-735, and NAGW-1474.

D. L. B. Jupp is with the Division of Water Resources, Commonwealth Scientific and Industrial Organization, P. O. Box 1666, Canberra, ACT, Australia.

A. H. Strahler and Curtis E. Woodcock are with the Department of Geography and Center for Remote Sensing, Boston University, Boston, MA 02215.

IEEE Log Number 8926437.

where Cov is defined in the usual way as the covariance between values distance h apart. The use of the variogram in geostatistics has been motivated by the existence of situations in mining geology where an infinite variance model is appropriate and by the fact that the local properties of the variogram (small h) are less affected by trends and regional patterns than the covariance.

Associated with the variograms of a regionalized variable is a local analysis called "Structural Analysis" [8], in which the basic properties of the variogram are used to assess the spatial structure of the regionalized variable. These are illustrated in Fig. 1. Note that this example is an isotropic variogram, in which the value of $V(h)$ is dependent only on the magnitude $|h|$ of h and not its direction. It can therefore be displayed as a simple line graph. In the case in which the variogram is not isotropic, and therefore depends on both the magnitude and direction of h, the variogram will have to be displayed as a three-dimensional surface through the use of contouring, perspective surface display, etc. [9].

There are five especially significant parameters that describe the variogram. These are:

1) The "sill," at which the variogram flattens off to a level equal to the general data variance. The existence and size of the sill are important parameters. The lack of a sill could indicate a trend in the data (e.g., a scan-angle brightness effect across the image) and/or that the domain of definition of the data is too small for the scales of pattern in the regionalized variable (i.e., h is never large enough to get to the sill).

2) The "range," the value of $|h|$ at which the variogram reaches the sill. This parameter is often related to the size or scale of the largest elements (objects) in the scene that give rise to the correlation structure.

3) The "nugget effect," which, when it occurs, is expressed as a finite limit for the variogram as $|h|$ tends to zero. The variogram is, of course, zero at $h = 0$. However, there may be uncorrelated noise in the observations that will always produce variance. In fact, if the regionalized variable is a realization of an uncorrelated noise series then the variogram will have no range and the nugget effect will equal the sill.

4) The "derivative at the origin," which is diagnostic of the essential variability of the data at the finest scale. If the limit of the derivative at the origin as $|h|$ tends to zero is near zero, then the underlying process will be "smooth"—whereas if it is high, then the process will be "rough." In our work this parameter can be interpreted as the amount of edge and boundary at the finest scale.

5) The "anisotropy" of the variogram, which can appear if the regionalized variable is defined in more than one dimension. It reveals a covariance structure such that values tend to be more similar in one or more preferred directions than in others. This will commonly arise in a remotely sensed image when the sun is not overhead and three-dimensional objects in the scene cast shadows (dark shapes) that are elongated in one direction [10], [11].

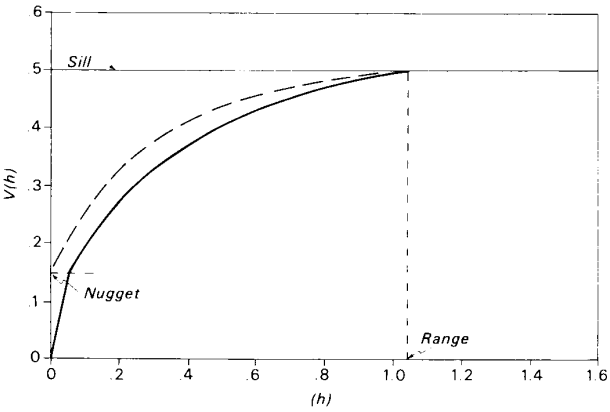


Fig. 1. Example of two variograms, showing sill, range, and nugget.

Cross-variograms can also be defined for multivariate spatial data—which is of special significance in remote sensing. In this case the sill is the variance/covariance matrix for the data, the nugget is the variance/covariance for the uncorrelated component in the image (usually regarded as noise), and the range will be a multivariate concept.

Among the parameters described above, the range is least well specified in practical terms. In many cases the model of a variogram has no specific point where it reaches the sill (such as for example an exponential model) and the presence of noise and sampling variance means that some other parameter is needed for estimating the range of observed variograms. Serra [12] defines a useful quantity related to the range

$$a_n(V) = \int_{R^n} \text{Cov}\,(h)\, d|h|/\text{Cov}\,(0) \qquad (3)$$

where R^n is an n-dimensional section of R^N, the full parameter space for the covariance. In the case of images, $n = 2$, and this concept of range is defined as an area in two dimensions or a set of one-dimensional measures defining range in specific directions. The behavior of a_1 with direction is clearly a measure of anisotropy, but for a two-dimensional isotropic covariance, a_1 and a_2 both represent a similar attribute—the linear range. In this case

$$a_1 = 2 \int_0^\infty \text{Cov}\,(r)\, dr/\text{Cov}\,(0) \qquad (4)$$

$$a_2 = 2\pi \int_0^\infty r\,\text{Cov}\,(r)\, dr/\text{Cov}\,(0). \qquad (5)$$

It is often used to convert a_2 to a linear measure by defining

$$a_2' = \left(\frac{a_2}{\pi}\right)^{1/2}. \qquad (6)$$

For example, for the exponential covariance $\text{Cov}\,(r) = \sigma^2 e^{-\alpha r}$ we find that $a_1 = 2/\alpha$ and $a_2' = \sqrt{2}/\alpha$, both of which provide direct estimates for α in the practical case.

II. Effects of Regularization on Spatial Statistics

A. Regularization and Matheron's Theorem

In practice, the variogram is computed from integrals over basic blocks, rather than on point measurements. For example, in remote sensing, the computed variogram for an image is based on integration of radiance over a field of view (pixels) and not the underlying radiance data. The underlying variogram is said to be "regularized" by this averaging over the shape of the pixel.

The term "regularize" relates to the situation in which a point process on the plane is convolved with a spatial filter supported on an element, or shape, $\mathbf{Z}$. If the points are regarded as having unit mass, and are thus the two-dimensional analog of the Dirac or delta function, the variance or variogram at small h will be ill-defined. However, if the points are convolved with the element $\mathbf{Z}$, the variable of interest becomes a count of the points within $\mathbf{Z}$. The new variable is now tractable—i.e., it is a "regular" function.

Suppose we have a regionalized variable $f(x)$ with variogram $V(h)$. Let $f_{\mathbf{Z}}(y)$ be the regularized function derived from $f(x)$ by convolving it with $\mathbf{Z}$. We can express this as

$$\begin{aligned} f_{\mathbf{Z}}(y) &= \frac{1}{\text{Mes}\,(\mathbf{Z})} \int_{\mathbf{Z}_y} f(x)\, d|x| \\ &= \frac{1}{\text{Mes}\,(\mathbf{Z})} \int_{\mathbf{R}^N} I_{\mathbf{Z}}(x + y)\, f(x)\, d|x| \\ &= \frac{f * \bar{I}_{\mathbf{Z}}(y)}{\text{Mes}\,(\mathbf{Z})} \end{aligned} \tag{7}$$

where $\mathbf{Z}_y$ is the regularizing element (such as a disk or a square) centered at y, Mes $(\mathbf{Z})$ is the measure of the element (the area, in the case $N = 2$), $I_{\mathbf{Z}}$ is the indicator function ($I_{\mathbf{Z}}(x) = 1$ for $x \in \mathbf{Z}$, 0 else), $\bar{I}_{\mathbf{Z}}(t) = I_{\mathbf{Z}}(-t)$, and $*$ is convolution. To avoid confusion, we will refer to the variogram of the underlying function ($V(h)$) as the punctual variogram and the variogram of the regularized function ($V_{\mathbf{Z}}(h)$) as the regularized variogram.

The results presented generally in Part I of this paper and the convolution form of Matheron's Theorem [3]–[5] can be used to relate the punctual variogram to the regularized variogram. In Part I, we showed how the effects of regularization on the covariance of a regionalized variable could be separated into the effects of the underlying covariance and the shape of the regularizing element. In the case where the regularizing process is the convolution above, the result reduces to the simple form

$$V_{\mathbf{Z}}(h) = (T * V)_h - (T * V)_0 \tag{8}$$

where T is the covariance (or overlap) function for the regularizing element $\mathbf{Z}$, or

$$T = I_{\mathbf{Z}} * \bar{I}_{\mathbf{Z}} / \text{Mes}^2\,(\mathbf{Z}). \tag{9}$$

Spelled out, the variogram of (say) radiance of pixels is related to the underlying variogram of radiance at points by a convolution of the underlying variogram with the covariance function of the pixel. This also means that the variogram for aggregations of pixels should be derivable for all aggregation block sizes if the punctual variogram is known. Note, however, that the usual case will be that the punctual variogram is not known and must be estimated from a regularized variogram.

B. Local Variance as a Function of Regularization

Consider the average variance between adjacent brightness values in a digital image of a discrete-object scene [1]—for example, the mean of a "texture" image, in which each brightness value is replaced by the standard deviation of brightness values within a three-by-three window centered around it. That is

$$T_{i,j} = \left[\tfrac{1}{8} \sum_{k=i-1}^{i+1} \sum_{l=j-1}^{j+1} (x_{k,l} - \bar{x}_{i,j})^2 \right]^{1/2} \tag{10}$$

where $T_{i,j}$ is the texture measure, and $\bar{x}_{i,j}$ is the mean within the three-by-three window around i, j. This type of variance has been referred to as the "local variance" by Woodcock and Strahler [13]. Let us now examine two cases. In the first, referred to as the H-resolution case [14], the pixels in the image are much smaller than the objects in the scene. Here, the local variance will be low, since many adjacent brightness measurements will have the same values—i.e., will take on either the brightness value of the objects or the brightness value of the background. In contrast is the L-resolution case, in which the pixels are much larger than the objects. In this situation, there will be many objects within each pixel, and if the density of objects is more or less uniform, then each pixel will tend to have about the same brightness. Thus, the local variance of such an image will also be low. Between these limits, however, there will be a peak in image variance where there is a maximum interaction between the pixel and object size [13].

In the context of this paper, the local variance is most properly expressed not as a "texture" value, but as the regularized value of the variogram at a step size equal to the regularization size. That is

$$V_{i,j} = \left[\tfrac{1}{16} \sum_{k=i-1}^{i+1} \sum_{l=j-1}^{j+1} |x_{k,l} - x_{i,j}|^2 \right]^{1/2}. \tag{11}$$

It is reasonable, too, to expect that the maximum or peak in a graph of local variance against block size will relate to the overall range of the underlying covariance function. In fact, in the practical case, the existence and position of the peak provide a more tractable analysis of the range than numerical integration of the estimated covariance function.

This can be seen by considering the case of an exponential model applied to the one-dimensional variogram, as in Part I of this paper. For this case, the punctual co-

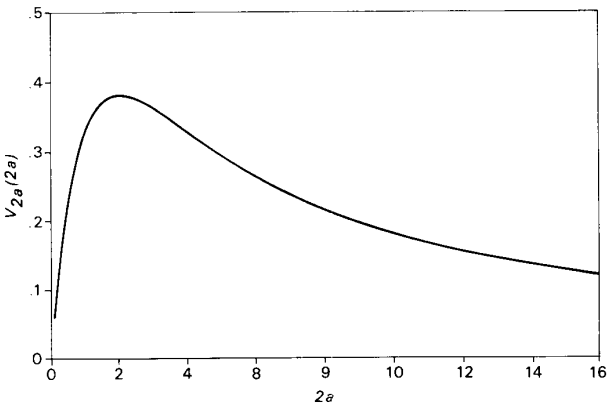


Fig. 2. Local variance (value of variogram at given block size) plotted as a function of block size $2a$ for a one-dimensional exponential model. Parameters σ^2 and α (see text) are taken as unity.

variance $C(h)$ with distance h is modeled as $\sigma^2 e^{-\alpha h}$. Expression (53) of that paper presents a formula for the variogram at step h regularized by a segment of length $2a$. Substituting $2a$ for h yields

$$V_{2a}(2a) = \sigma_{2a}^2 - \frac{\sigma^2}{(2\alpha a)^2}(1 - e^{-2\alpha a})^2. \quad (12)$$

From this expression, it can be shown that the peak will be reached at $2a = 1.8926/\alpha$, and will have a value of $0.3811\sigma^2$. Fig. 2 shows a graph of this local variance plotted as a function of block size. Finally, it should be noted that this analysis is very similar in concept, and in some cases identical, to the methods based on hierarchical analysis of variance used to study ecological and geographic patterns by, among others, Grieg-Smith [15], Hill [16], and Moellering and Tobler [17].

III. Images Derived from Disk Models

In the usual geostatistical applications of regionalized variables, it is common to assume an arbitrary covariance structure such that the covariance reduces with distance in (e.g.) a linear, exponential, or spherical fashion. The first paper in this series [2] provided some examples of exponential covariance structures, regularized in one and two dimensions by line segments and disks, respectively. However, when we consider the case of a discrete-object model for a remotely sensed scene, it may be possible to derive an explicit covariance function based on the sizes, shapes, brightnesses, and densities of objects. Perhaps the simplest such scene model of this nature is that of overlapping disks of constant size and brightness that are randomly placed on a contrasting background. If forests or shrublands are regarded as land surfaces covered by objects that have more or less round projections onto the ground, it is easy to see that this simple model may be relevant to a real application. Further, the random disk model is numerically tractable, utilizing simple, isotropic geometry and well-known theorems from the field of geometrical probability. And, it is not overly difficult to extend the model to populations of disks with varying size (e.g., lognormal) distributions or to clustered (e.g., Neyman Type A) location functions, as we will show in a subsequent section.

A. The Simple Disk Model

The first model considered is the simple one of disks of constant radius scattered randomly on the plane (Fig. 3). The radiances (or "colors") of the disk and background are taken as 1 and 0, respectively. This is no limitation as long as it is assumed that overlapping disks have the same radiance as single disks. If this simple "scene" is imaged such that the brightness value of a pixel represents the radiance averaged over an area $\boldsymbol{A}$ of the pixel, then the data are simply the fraction of each resolution cell covered by disks. The model results in a random set function that is clearly stationary and isotropic in the sense defined in the first paper of this series. The effect of integration of the scene into image pixels is to regularize the underlying regionalized variable.

From our work in the first paper, it follows that the basic set quantities needed to construct the punctual variogram for this scene model are the proportion of the uncovered area (or the probability Q_1 that a point x will fall on the background), which is

$$Q_1 = \text{Prob}\{x \in \boldsymbol{D}^c\} \quad (13)$$

and the set overlapping function of the uncovered area Q_2 (or the probability that a point x and its neighbor distance h away will both be in the background)

$$Q_2 = \text{Prob}\{x \in \boldsymbol{D}^c \cap \boldsymbol{D}^c_{-h}\} \quad (14)$$

where $\boldsymbol{D}$ is the area covered by disks and $\boldsymbol{D}^c$ is the complement of this area—i.e., the area of background not obscured by the disks—and $\boldsymbol{D}^c_{-h}$ is the unobscured background shifted by $-h$. If these are known, the covariance function for points at distance h in $\boldsymbol{D}^c$ is

$$\text{Cov}(h) = Q_2 - Q_1^2. \quad (15)$$

When the centers of the disks are distributed as a Poisson function, it follows that the expectations for these quantities are respectively [12]

$$Q_1 = e^{-\lambda A} \quad (16)$$

and

$$Q_2 = Q_1^2 e^{\lambda K(h)} = Q_1^2 e^{\lambda A T(s)} \quad (17)$$

where a disk has area $\boldsymbol{A}$ and diameter D, and λ is the density of disk centers in counts per square areal unit. $K(h)$ is the set of overlap function (area of overlap) between individual disks as a function of the distance h between them, which may be expressed in terms of the scale-free function $T(s)$ as $K(h) = \boldsymbol{A}T(s)$, where s is the normalized distance h/D. For the disk shape

$$T(s) = \frac{\theta - \sin\theta}{\pi} \quad (18)$$

where $\cos(\theta/2) = s$.

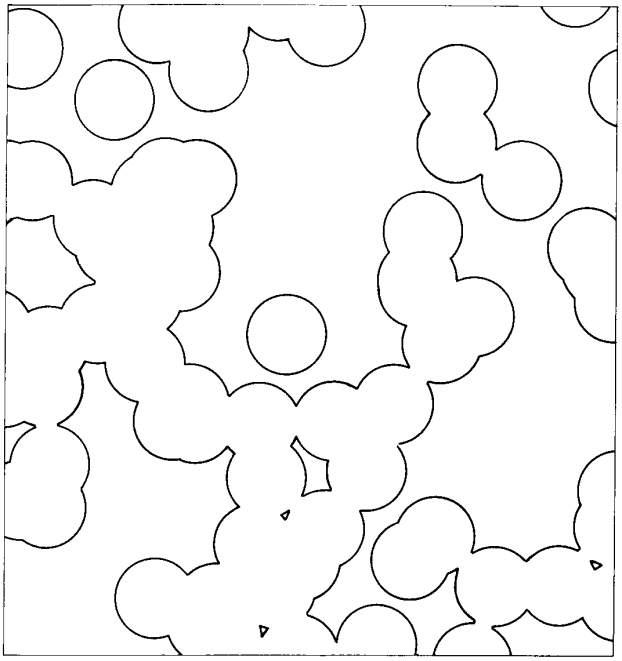


Fig. 3. Discrete-object scene of overlapping disks placed randomly on a background. Cover is 50 percent.

From these expectations, the covariance function for the background of the disk model can be determined:

$$\begin{aligned} \text{Cov}\,(h) &= Q_1^2(e^{\lambda K(h)} - 1) \\ &= Q_1^2(e^{\lambda A T(s)} - 1). \end{aligned} \tag{19}$$

The variogram is equally well-defined and has the form

$$\gamma(h) = Q_1^2[1/Q_1 - e^{\lambda A T(s)}] \tag{20}$$

and is the same for disks and background. The punctual disk variogram has the following five properties:

1) The global variance, or "sill," is $Q_1(1 - Q_1)$, which is the (binomial) variance for the random set function.

2) The variogram reaches the sill at $h = D$, the (constant) diameter of the disks.

3) The variogram passes through the origin, i.e., does not exhibit a nugget effect.

4) The derivative of the variogram at the origin is $\gamma'(h) = \lambda A e^{-\lambda A}$. Structurally, this is proportional to the effective amount of boundary between the disks and the background and is maximized when the cover is about 63.2 percent. (This point of "maximal patchiness" is different from the cover at which variance is maximum—i.e., 50 percent.)

5) The variogram is isotropic, which arises because the overlap function is isotropic.

Properties 2) and 4) also imply that although neither the mean nor variance of the radiance nor the derivative at the origin of the variogram for the underlying model is sensitive separately to disk size and spacing, the distance to the sill *is* sensitive, and appears even at this stage as an important diagnostic statistic.

The linear range (a_1), discussed in a preceding section, has the form

$$a_1 = D\,\frac{2Q_1}{1 - Q_1}\int_0^1 (e^{AT(s)} - 1)\,ds \tag{21}$$

and can either be computed numerically or approximated using the substitution $T(s) \simeq 1 - s$ in the above integral, in which case we find that

$$a_1 = 2D\,\frac{Q_1}{1 - Q_1}\left[\frac{1 - Q_1}{AQ_1} - 1\right] \tag{22}$$

or

$$\frac{a_1}{D} = F(Q_1) \tag{23}$$

which serves to show a_1 as a separable function of Q_1 and D. For small densities, a_1 approaches $2D$, or twice the disk diameter. In the practical case, therefore, the linear range a_1 has the same diagnostic use as the (exact) distance to the sill. In the case of actual data and regularized models, this parameter will assume greater significance.

B. The Regularized Disk Model

It remains to regularize the image of disks by a pixel of fixed shape. In remote sensing, we typically regard the pixels in the image as being of square or rectangular shape. This is largely for the convenience of apportioning the image space to values associated with a regularly spaced grid. In truth, the instantaneous field of view of a detector will be blurred by optical and electronic effects, even when the detector surface may be rectangular. Thus, a circular shape for regularization is not inappropriate, and we will proceed with a scene model of disks of diameter D_1 imaged by a circular pixel of diameter D_2.

We will consider first the mean, then the covariance, of the disk model regularized by a disk. Since the model is spatially stationary, the expected value of the datum in each pixel is the same as the expected cover ($1 - Q_1$) at any point on the plane. This follows since

$$Q_1 = \frac{1}{A}\int_A \text{Prob}\,\{x \in D^c\}\,d|x|. \tag{24}$$

That is, the expectation that a point is in the pixel and hits a disk is the same as the expectation that any single point hits a disk—and equals $1 - Q_1$.

Since the disk model and the regularizing set are isotropic, the general covariance formulae of above and our earlier paper [2, equations (43) and (47)] may be applied. That is, the variogram for the regularized disk is

$$\gamma_Z(s) = \sigma_Z^2 - \text{Cov}_Z\,(s) \tag{25}$$

where $s = h/D$ and

$$\sigma_Z^2 = 8\int_0^1 tT(t)\,\text{Cov}\,(Dt)\,dt \tag{26}$$

$$\mathrm{Cov}_Z(t) = \frac{8}{\pi} \int_{\max(0,-1+s)}^{1+s} t\,\Phi(t, s)\,C(Dt)\,dt. \quad (27)$$

The variance (sill of the variogram) is σ_Z^2 and the local variance, or variogram at step size equal to the diameter of the regularizing disk ($h = D$, $s = 1$), is

$$\gamma_Z(1) = \sigma_Z^2 - \mathrm{Cov}_Z(1). \quad (28)$$

The effects due to image regularization on the punctual variogram are shown in Fig. 4, which presents the result of regularizing a disk model with a 50-percent cover and basic disk diameter of one unit with disk-shaped pixels ranging in size from one to four units. As the image is regularized by successively larger pixels, the total variance (or sill) decreases, the range increases (to $D_1 + D_2$, where D_1 and D_2 are the diameters of the basic and regularizing disk), and the derivative at the origin becomes zero.

Fig. 5 presents variograms for scenes of varying disk size as regularized by a disk of unit diameter. The figure displays the change in the variogram as the scene model goes from L-resolution ($D_1 \ll D_2$) to H-resolution ($D_1 \gg D_2$). As above, the cover for each scene model is 50 percent.

Fig. 6 plots the local variance (value of the variogram at step equal to the regularizing element size) as a function of D_2/D_1. The figure shows a well-developed peak at a ratio of D_2/D_1 equal to about 0.7. Both the size of the peak and the precise position depend on the cover fraction, but for a given cover ($1 - Q_1$) there will be a unique ratio D_2/D_1 at which it occurs. Thus the position of the peak can reveal the size of the underlying disks and provides a readily computible and practical method in the L-resolution case.

The tail end of Fig. 6 approaches the regularized overall variance C_Z asymptotically, which also contains information on the scaling properties of the image. This can be seen in Fig. 7, which plots log (C_Z) against log (D_2/D_1) and shows the overall effects of regularizing the variance. The plot falls slowly at first, and then approaches -1 asymptotically as would be expected if the image were purely random. Also plotted is a line with slope -0.75, which is a slope noted by H. Fairfield Smith [18] for similar plots involving yields of agricultural crops. These effects will be discussed further later.

C. The Effect of Subpixel Structure—L-Resolution Models

The peak of the local variance against block size is diagnostic for the size of the disks. However, it occurs for a regularizing disk with a diameter D_2 that is slightly smaller than the size of the underlying disk with diameter D_1. This is then the H-resolution case described above. For variable-sized distributions of objects, the smallest objects in the image will behave as for an L-resolution scene model and induce a behavior in the local and global variance that may be indistinguishable from that due to a random point process.

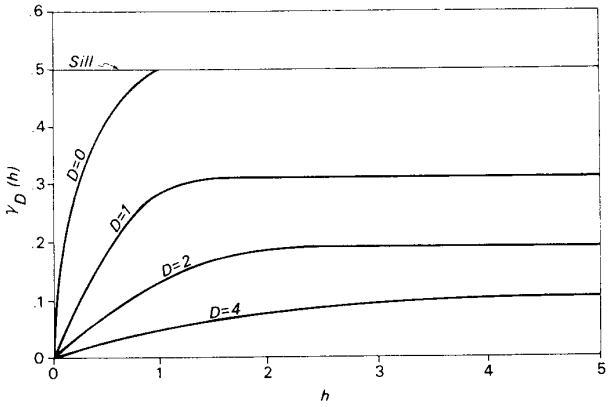


Fig. 4. Regularized variograms for a scene of overlapping disks of unit diameter covering 50 percent of the background. D: diameter of regularizing disk. $D = 0$ denotes the unregularized (punctual) variogram.

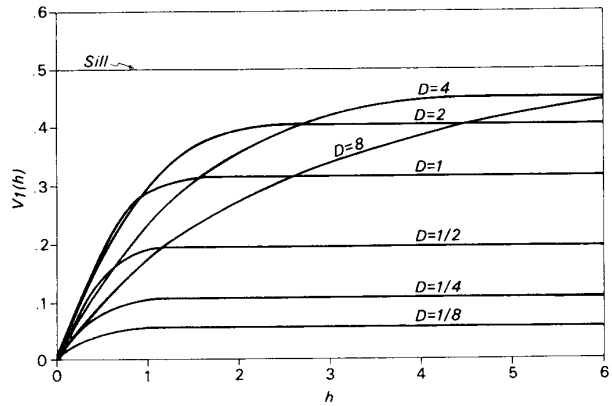


Fig. 5. Variograms for disk-model scenes with disk diameters ranging from 0.0625 to 8 units, regularized by a disk of unit radius. Cover is 50 percent for all cases.

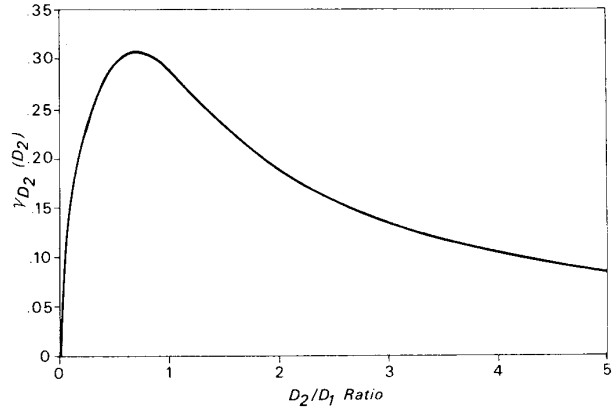


Fig. 6. Local variances (covariance at step size D_2) as a function of D_2/D_1 for the regularized disk scene.

As noted earlier, the indicator function for a point process is a delta function, which does not lead to a definable covariance. However, if the point process is regularized

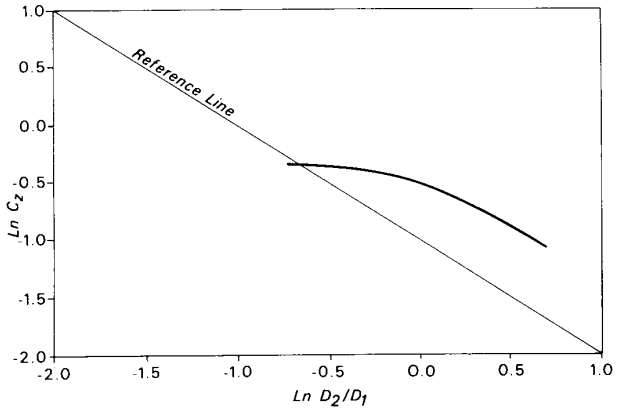


Fig. 7. Log of the regularized covariance of the disk model plotted against the log of the ratio D_2/D_1. The reference line has slope -0.75.

by a small disk (diameter D_1), then the random function becomes quite well-behaved, being the number of points within the radius of the disk at each point of the plane. The resulting process is similar to, but not the same as, a disk model with disk diameter equal to that of the regularizing disk D_1. They have the same backgrounds and therefore the same set covariances. However, in the regularized point process the areas where disks overlap take different "radiance" values from areas of disk where there is no overlap. The two become more alike, however, as the regularizing disk size decreases and if both are further regularized into a "pixel model"—such as a disk of diameter D_2.

The variogram for small disk sizes and a given pixel size is similar in form to that of a large pixel for given disks—the form depends only on the ratio D_2/D_1. In both cases the model is generally indistinguishable from a regularized point process. The variogram at a step of one pixel is likely to be nearly identical with the sill and they both decrease monotonically with the area of the regularizing pixel.

This effect is most significant when there is a distribution of disk sizes. Then those disks much smaller than the pixel will contribute to the variogram in a similar way to a point process, producing a nugget effect, while the large disks will contribute like the H-resolution case.

IV. Generalizations of the Disk Model

The simple disk model, with constant size and Poisson spacing, contains much of the structure needed to relate variograms of scenes and images in a meaningful way. However, it is clearly very limited. In this section, we extend the disk model to accommodate disks of varying size as well as spacing functions that are alternative to the Poisson.

A. Derivation of the General Model

To generalize the model, we will return to first principles and derive a model for punctual scene covariance that is quite general. Note that for the Poisson spacing model, it was possible to neglect the extent of the scene in which the model occurred and thus ignore edge effects. However, our more general model will need to include considerations of this problem, and thus it must be posed as follows.

Using the same assumptions as Garwood [19], let $\boldsymbol{A}$ be a particular realization of a collection of random discrete objects with area A in the plane, and $\boldsymbol{S}$ be the scene with area S in which the objects are located within a larger field $\boldsymbol{S}'$. Each object has a particular point that acts as its "origin," and the general model sees objects $\boldsymbol{A}$ scattered randomly in $\boldsymbol{S}$ with origins located according to some point process in a subset of $\boldsymbol{S}$ so that the objects fall totally in $\boldsymbol{S}$ (Fig. 8).

For now, consider that only one object from the collection is located with an origin in $\boldsymbol{S}$. Then, over many realizations of $\boldsymbol{A}$ we will find

$$\text{Prob}\left\{x \in \boldsymbol{A}^c\right\} = E\left\{\frac{S - A}{S}\right\} = \frac{S - \bar{A}}{S} = q_1 \tag{29}$$

$$\text{Prob}\left\{x \in \boldsymbol{A}^c \cap \boldsymbol{A}^c_{-h}\right\} = E\left\{\frac{S - 2A + K(h)}{S}\right\} = \frac{S - 2\bar{A} + \bar{K}(h)}{S} = q_2 \tag{30}$$

where $K(h)$ is the overlap function between a realization of $\boldsymbol{A}$ and itself and $\bar{K}$ is the average of K over the realizations. For disk models, for example, if the probability distribution function for the disk diameter is $p(D)$, this means that

$$\bar{A} = \frac{\pi}{4}\int_0^\infty D^2 p(D)\, dr \tag{31}$$

and

$$\bar{K}(h) = \int_0^\infty K(D, h)\, p(D)\, dr. \tag{32}$$

Now consider N objects independently realized in $\boldsymbol{S}$. In this case, $Q = q^N$, and if the probability of N is specified as $p(N)$, then

$$Q = G(q) = \sum_{N=0}^{\infty} p(N)\, q^N \tag{33}$$

which can be regarded as the generating function $G(u)$ for the point process. For the Poisson distribution, this generating function becomes

$$G(u) = \sum_{N=0}^{\infty} \frac{e^{-\lambda S}(\lambda S)^u}{u!} u^N = e^{\lambda S(u-1)} \tag{34}$$

where λ is a density parameter in counts per unit area, and λS is thus the mean number of objects with S. From

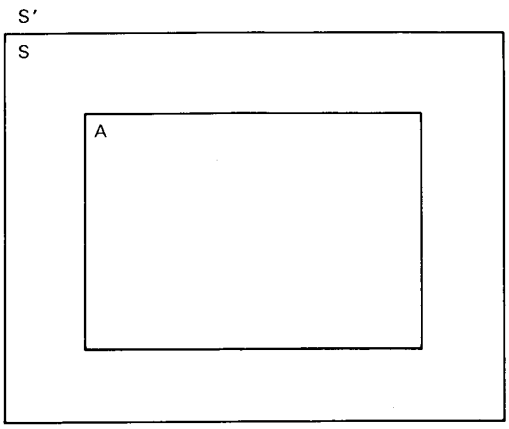


Fig. 8. Diagram of space S' showing S and A within it.

the earlier definitions in this section

$$S(q_1 - 1) = -\bar{A}$$
$$S(q_2 - 1) = -2\bar{A} + \bar{K}(h) \tag{35}$$

and thus

$$Q_1 = e^{-\lambda\bar{A}} \tag{36}$$
$$Q_2 = e^{-\lambda(2\bar{A} - \bar{K}(h))} \tag{37}$$

from which Cov (h) $(= Q_2 - Q_1^2)$ is easily derived. This result is identical with the previous simple model when the objects are disks of constant size and similar derivations could be used with other common point processes such as geometric, binomial, or negative binomial.

B. Clustered Models

To obtain effects due to clustered models, consider the Neyman Type A point process model as an example. This distribution has the generating function

$$G(u) = e^{-\mu C(1 - E(u))} \tag{38}$$

where

$$E(u) = e^{\mu(u-1)/C} \tag{39}$$

and C is the mean area of the clusters. However, the covariance does not result immediately from substituting q_1 and q_2 above into $G(u)$ since the clusters and the individuals making them up are not independently defined. To derive formulae for the overlap function for clustered models, we need to consider how the clustering process is being defined. For example, a Neyman Type A distribution will occur if the following two-stage model is used:

> The centers of a "parent" process are Poisson distributed with density μ in S. Associated with each parent is an area bounded by a disk with area A_c. Within each disk the centers of the objects, with a mean area A_d, are distributed with Poisson density m.

Now the problem can be approached with the clusters as the new set of objects and the complete result constructed as above.

With some manipulation, it follows that

$$A = A_c(1 - q_1) \tag{40}$$

where $q_1 = e^{-mA_d}$ and

$$K(h) = K_c(h)\,[1 - 2q_1 + q_1^2 e^{mK_d(h)}] \tag{41}$$

where K_c is the overlap function for the areas associated with the parent process and K_d is the overlap function for the objects within these areas. Hence, for this construction of a clustered model

$$Q_1 = e^{-\mu A_c(1 - q_1)}$$
$$\text{Cov}\,(h) = Q_1^2(e^{\mu K(h)} - 1). \tag{42}$$

It is clear that such a model will exhibit two scales of pattern relative to the sizes of the objects (e.g., "trees") and clusters (e.g., "patches") in a way that can be demonstrated both from the underlying covariance and from the effects of regularization.

In the more general case than the Neyman, the objects associated with each cluster could be located a distance r away from its center according to some probability law $P(r)$, or other distributions could be used for the parent process. Some of these are listed in Thompson [20].

C. Effects of Generalization on Variogram Statistics

The model generalization introduces some significant effects into the variograms and local variance statistics and provides a guide to the kind of parameters that may be accessible by model inversion. As well, generalization provides a basis for the observed statistics of real images. First, a reasonable method for the disk diameters is the lognormal model

$$p(D) = \frac{1}{D\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left[\frac{\log(D) - \mu}{\sigma}\right]^2\right\} \tag{43}$$

where the log mean (μ) and the log variance (σ^2) are related to the mean and variance of the disk diameters as $\bar{D} = e^{\mu + \sigma^2/2}$ and $s_D^2 = e^{2\mu + 2\sigma^2}$. In this case it follows that

$$\bar{A} = \frac{\pi}{4}\,[\bar{D}^2 + s_D^2] \tag{44}$$

and $\bar{K}(h)$ needs to be computed numerically.

It is clear in this case that a lognormal disk model and a fixed size model with an effective disk diameter of

$$D_{\text{eff}} = \bar{D}(1 + CV^2)^{1/2} \tag{45}$$

where CV is the coefficient of variation of the distribution, will have the same density (λ) if they have the same cover fraction. This follows since

$$1 - Q_1 = e^{-\lambda\bar{A}} \tag{46}$$

in each case.

Fig. 9 shows the effect of variable disk sizes by plotting a series of punctual variograms corresponding to the in-

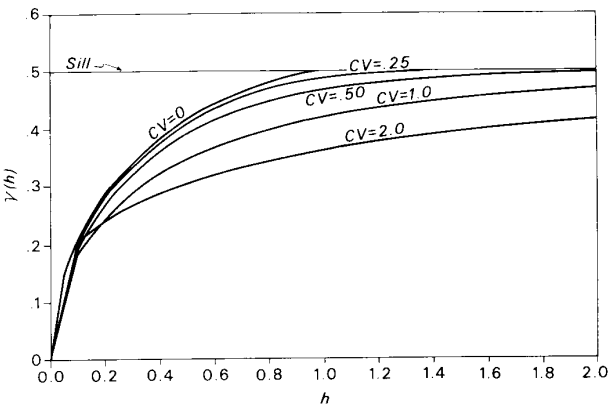


Fig. 9. Punctual variograms for disk models with lognormally distributed disk radii.

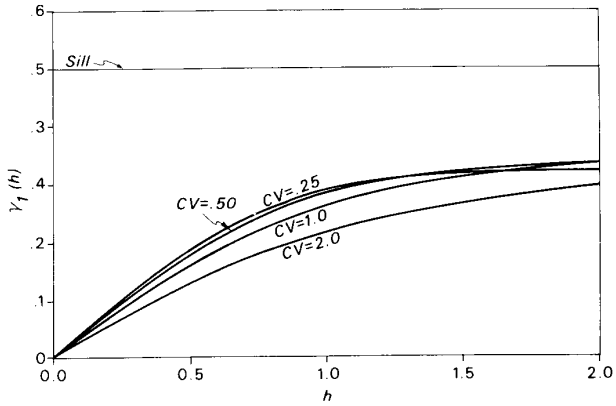


Fig. 10. Variograms for disk models with lognormally distributed disk radii as regularized by a disk of unit radius. *CV*: Coefficient of variation.

creasing coefficient of variation but with the mean radius chosen such that all variograms correspond to disk models with the same cover fraction and the same effective fixed disk radius. Fig. 10 shows the same result but with regularization by a disk of unit diameter, and Fig. 11 shows the effect on the local variance of the increasing coefficient of variation. It is clear that for a coefficient of variation greater than one, the effects of the diameter distribution are significant in both the regularized variogram and the variance/block size diagram. The difference is enough to suggest that the inversion of mean size, spacing, and *CV* from image models is well posed if the complete variogram function is used. For high *CV*, the influence of the large disks shows in the extended sill and peak of the local variance, and the small disks combine to produce a nugget effect in the regularized variogram. However, it should be noted that a *CV* near 0.5 would be typical of most open forest stands.

To investigate the effects of clustering, we computed variograms for a model where "tree crowns" (i.e., disks) of 1 m diameter were clustered with cluster diameters of 9 m such that the overall cover was 50 percent and the

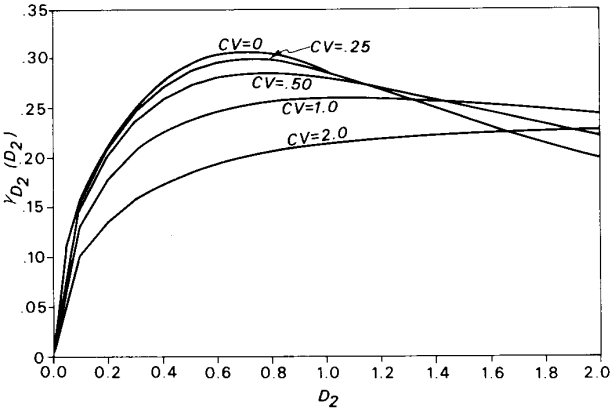


Fig. 11. Local variance as a function of block size for models of disks with lognormally distributed radii.

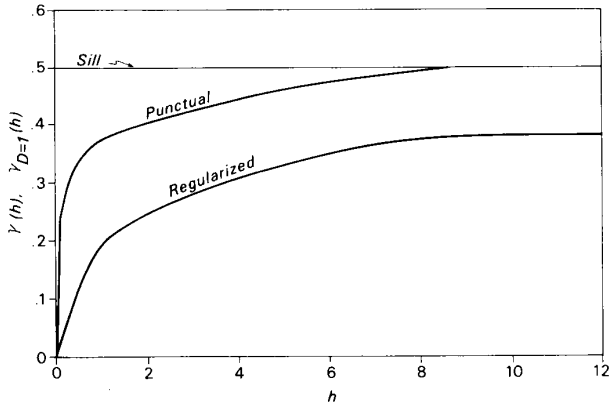


Fig. 12. Punctual and regularized variograms for disks of clustered spacing. Diameter of regularizing disk = 1.

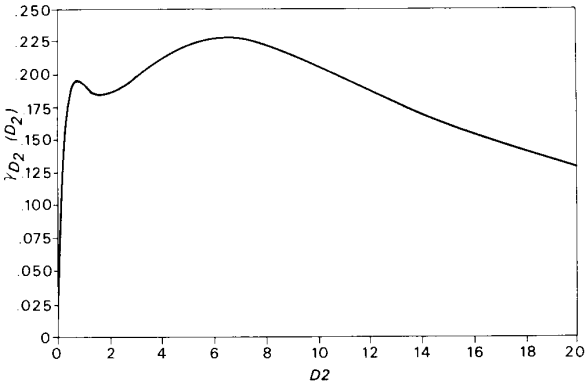


Fig. 13. Local variance plotted against block size for the Neyman Type A clustered model.

within cluster cover was 60 percent. The resulting punctual and regularized variograms are plotted in Fig. 12. Note that even though there are two scales of pattern in the scene model, that fact is not obvious in Fig. 12.

Fig. 13 is a plot of local variance against regularization in which two obvious peaks corresponding to the two

scales of pattern are clearly visible. These effects are masked in the variograms, but appear visibly in the local variance plot. This suggests that the variogram should always be evaluated in the light of a preliminary analysis of the effects of blocking, or increased regularization, on both the local and overall image variance.

V. An Application to Plot Size and Cover Estimation

The general relationship between cover fraction, image variance, and regularization can be illustrated by the following application of the methods derived in this paper. This example shows how the plot size for field vegetation sampling may be determined from the expected spatial covariance for random disks.

The transition from an H-resolution model, in which the tree crowns are either visible or near resolvable directly, to the L-resolution model, in which the crowns are subpixel effects, represents a loss of information for the image as a whole but a gain at the pixel level. For example, in field data collection, cover and canopy structure is estimated over a plot which is a kind of field "pixel." It is usually assumed that the estimate is at least locally stable in the sense that taking another plot nearby would not give dramatically different results. Again, much classificatory work on images consists of estimating the cover or condition at a single pixel level. It is clear that such an estimate will improve as the pixel size increases up to a point at which different covers are being integrated into the pixel.

Consider a simple example, in which the crowns are all 8 m in diameter with Poisson density. As described above, the radiance of a pixel will be proportional to the fraction of crown P or

$$P = \frac{A_c}{A} \tag{47}$$

and has an expectation $1 - Q_1$, the cover.

If we wish the estimate P in a plot to be stable in the sense that the cover of an adjacent plot will fall into the same cover class, then this translates into the formulae of this paper as saying that the local variance should be small enough that the adjacent pixel or plot will be of the same type. For field data collection, this stability is important, although total variance (but not local variance) can be kept by taking data at a number of sites. In image analysis, cover or class estimates will "spot" from pixel to pixel if the pixel is too small for the spatial variation.

To define some criteria, we can use standard cover classes for forested areas of the form:

Cover Type	Percent
Forest	80–100
Open Forest	50–80
Woodland	20–50
Open Woodland	0–20

TABLE II
Expected Range (±2 Standard Deviations) in Adjacent Plot Cover (Percent)

Pixel Size	Cover Type and Percentage			
	Forest (80-100)	Open Forest (50-80)	Woodland (20-50)	Open Woodland (0-20)
10 m	65-100	17-100	0-85	0-44
20 m	75-100	37-93	3-67	0-34
30 m	80-100	45-85	13-57	0-24
50 m	84-96	53-77	21-49	1-19
80 m	86-94	57-73	26-44	4-16
480 m	89.3-90.7	63.6-66.4	33.6-36.4	9 - 11
800 m	89.7-90.3	64.3-65.7	34.2-35.8	9.5-10.5
1.1 km	89.8-90.2	64.5-65.5	34.4-35.6	9.6-10.4

and see to what degree adjacent plots (or pixels), represented by a mid-range value, will be allocated to the same class as a function of plot size. To give the plot sizes, some context for remote sensing the sizes will match the resolutions of the main satellite sensors:

Sensor	Resolution
NOAA AVHRR	1.1 Km
NIMBUS CZCS	800 m
HCMM	480 m
Landsat MSS	80 m
MOS-1	50 m
Landsat TM	30 m
SPOT (MSS)	20 m
SPOT (Panchromatic)	10 m.

Table II shows the limits based on two standard deviations (5 percent level) about the mid-range value in each class. It is clear for this simple model that it is only at a plot size of 50 m that if one plot has the mid-range value of a class, its neighbor has a 90 percent chance of being in the same class. In patchy open forests and woodlands, therefore, plot sizes up to a 100 m diameter may be needed. Clearly, if also the pixels fall into consistent cover classes, the sensors with pixel sizes in the HCMM/CZCS/AVHRR range provide very stable cover estimates—but almost no subpixel texture information.

These results show how the covariance at the finest scale can be used to investigate plot size, but if there is clustering of the form modeled above, there a further investigation is needed of the spacing of the plots to adequately sample the pattern.

VI. Discussion

A. The Essential Properties of Disk Models

In remote sensing, the data are scene radiance integrated over finite areas (pixels), sampled at the pixel spacing and collected into a complete image. From our previous work it follows that for objects with radiance 1 on a background of radiance 0, the expectation of the pixel radiance is just $1 - Q_1$, the cover fraction,

$$1 - Q_1 = 1 - e^{-\lambda \bar{A}}.$$

This formula is especially significant in its implications for remote sensing of land cover structure, since the density and size of disks are not separable from mean radiance—which reflects only the total object area per unit area $\lambda\bar{A}$. For example, if the disks are identified with the vertical projection of trees or shrubs, then this result suggests that pixel radiance is insensitive to whether it is being generated by a few large plants or many small ones, provided that in each case the cover is the same. However, this is not the case for the second moment, for which pixel size interacts with the size of the disks. Li and Strahler [21], [22] have used this fact in their inversion of conifer forest models and it may represent the most immediately significant opportunity to be had from covariance or variogram analysis in remote sensing.

For models with a restricted number of parameters (such as mean tree height and spacing, with the height/diameter ratio and the coefficient of variation of a lognormal height distribution fixed), Li and Strahler used the radiance data to estimate the radiances of the components, and therefore Q_1. They then estimated the sizes of the objects by the relation between the sill (total variance), the degree of regularization, and the object size. It is clear that the structure of the complete variogram described here provides an avenue for further development of their methods, especially when multicomponent systems with shadowing are being considered [23].

Previous work [10], [11] has indicated that the variograms are not very sensitive to the shapes of the pixels and objects, provided that they are not anisotropic. The results presented here also show (using the lognormal model) that if the variance of size is not too high, it is again a marginal effect in the data. Overall then, the dominating effects are the average size and spacing of the objects, which show themselves in the relationship between the mean radiance and variance of radiance. It is quite likely, for example, that in the L-resolution case these are the only parameters available from the second-order spatial statistics.

When the pixels provide an H-resolution model for at least some of the scene features (clusters of smaller objects, for example), the graph of local variances against pixel size can provide a sensitive diagnostic to the scales of pattern in the image. It seems that this graph is one of the best tools for the initial investigation of spatial pattern and variance. The eventual rate at which this statistic decreases for large blocking is also an indicator of whether the scene is large enough to encompass the basic variation being imaged or whether it is part of a larger pattern.

B. Scaling Properties of the Models

The scaling effects for the disk models in the global sense can be investigated using the plot of log C_Z against log D_2/D_1 as in a previous section. Fig. 14 contains the results for a fixed size disk, for lognormally distributed sizes with a coefficient of variation of 1.0, and for the clustered model. Also plotted are a line with slope -1.0 and one with slope -0.75. These lines represent asymptotes for essentially random data and the "fractal" behavior noted by H. Fairfield Smith [18]. From the plots, it is clear that all of the models eventually behave as random data with the clustered model retaining significant nonrandom behavior for the greatest range. This plot, then, gives us a clue to the scaling behavior found in most real scenes [10], [11], [13] in which there exists a hierarchy of scales of pattern that any effective land surface model must address. For example, leaves cluster into crowns of trees, trees into patches, patches into forests or woodlands, and so on. The behavior this induces is sometimes called "fractal," and the methods outlined in this paper provide a means for studying some of the sources of fractal behavior over a range of scales.

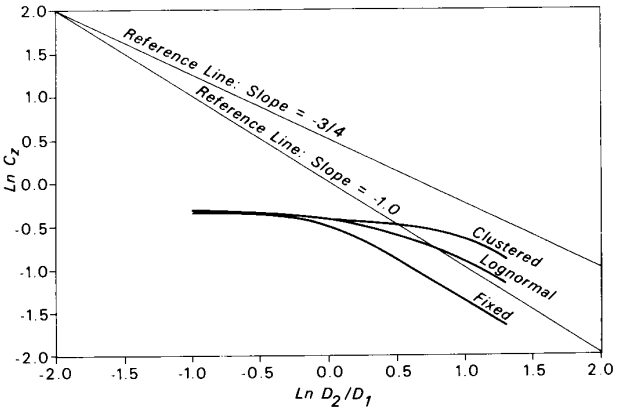


Fig. 14. Plot of the log of regularized covariance (C_Z) with D_2/D_1 ratio for clustered, lognormal, and fixed-size disk models.

VII. Conclusion

The methods of spatial analysis outlined here are sufficient to provide significant definition of the structure of simple scene models. This is shown by our analysis of the disk models, including effects of regularization, clustering, and varying size distribution. Despite their simplified nature, the disk models demonstrate most of the features and analytical methods that are now being used to model the cover and structure in images of complex, shadowed three-dimensional scenes. The clustered model already includes many of the tools needed to model tree crowns as clusters of leaves, or patches of forest as clusters of trees, and the overlap function methods have been used to describe the bidirectional reflectance distribution function [22] and hotspot [23] effects that the combination of canopy structure, view angle, and sun angle induce in a landscape. Thus, the methods should be immediately generalizable to real scenes in real remote sensing situations [10], [11].

There remains for future work 1) the modeling of spatial variance properties of more complex scenes, made up of three-dimensional objects and their shadows, and 2) the further development of inversion methods that directly reveal the characteristics of the underlying population of discrete objects, especially in the L-resolution case.

REFERENCES

[1] A. H. Strahler, X. Li, C. E. Woodcock, and D. L. B. Jupp, "Discrete-object modeling of remotely sensed scenes," in *Proc. 19th Int. Symp. Remote Sensing Environ.* (Paris, France), Oct. 1-5, 1984, pp. 465-473, 1985.

[2] D. L. B. Jupp, A. H. Strahler, and C. E. Woodcock, "Autocorrelation and regularization in digital images I. Basic Theory," *IEEE Trans. Geosci. Remote Sensing*, vol. 26, no. 4, pp. 463-473, July 1988.

[3] G. Matheron, "Principles of geostatistics," *Economic Geology*, vol. 58, pp. 1246-1266, 1963.

[4] ——, *Les Variables Regionalisées et Leur Estimation*. Paris: Masson et Cie., 1965.

[5] ——, "The theory of regionalized variables and its applications," in *Cahiers du Centre de Morphologie Mathematique de Fontainebleau*, no. 5, 1971.

[6] L. S. Gandin, *Objective Analysis of Meteorological Fields*. Springfield, VA, translated from Russian by Israel Program for Scientific Translations, 1963.

[7] G. H. Jowett, "Sampling properties of local statistics in stationary stochastic series," *Biometrika*, vol. 42, pp. 160-169, 1955.

[8] C. J. Huijbregts, "Regionalized variables and quantitative analysis of spatial data," in *Display and Analysis of Spatial Data*, J. C. Davis and M. J. McCullagh, Eds. London: Wiley, 1975, pp. 38-53.

[9] C. E. Woodcock, "Understanding spatial variation in remotely sensed imagery," Ph.D. dissertation, Univ. of California, Santa Barbara, 1985.

[10] C. E. Woodcock, A. H. Strahler, and D. L. B. Jupp, "The use of variograms in remote sensing I: Scene models and simulated images," *Remote Sensing Environ.*, vol. 25, pp. 323-348, 1988.

[11] ——, "The use of variograms in remote sensing II: Real digital images," *Remote Sensing Environ.*, vol. 25, pp. 349-379, 1988.

[12] J. Serra, *Image Analysis and Mathematical Morphology*. London: Academic, 1982.

[13] C. E. Woodcock and A. H. Strahler, "The factor of scale in remote sensing," *Remote Sensing Environ.*, vol. 21, pp. 311-322, 1987.

[14] A. H. Strahler, C. E. Woodcock, and J. A. Smith, "On the nature of models in remote sensing," *Remote Sensing Environ.*, vol. 20, pp. 121-139, 1986.

[15] P. Grieg-Smith, "The use of random and contiguous quadrats in the study of the structure of plant communities," *Ann. Bot., N. S.*, vol. 16, pp. 293-316, 1952.

[16] M. O. Hill, "The intensity of spatial pattern in plant communities," *J. Ecol.*, vol. 61, pp. 225-235, 1973.

[17] H. Moellering and W. Tobler, "Geographical variances," *Geog. Anal.*, vol. 4, pp. 35-50, 1972.

[18] H. Fairfield Smith, "An empirical law describing the heterogeneity in the yields of agricultural crops," *J. Agric. Sci.*, vol. 28, pp. 1-23, 1938.

[19] F. Garwood, "The variance of the overlap of geometrical figures with reference to a bombing problem," *Biometrika*, vol. 34, pp. 1-17, 1947.

[20] H. R. Thompson, "A note on contagious distributions," *Biometrika*, vol. 41, pp. 268-271, 1954.

[21] X. Li and A. H. Strahler, "Geometric-optical modeling of a conifer forest canopy," *IEEE Trans. Geosci. Remote Sensing*, vol. GE-23, no. 5, pp. 705-721, Sept. 1985.

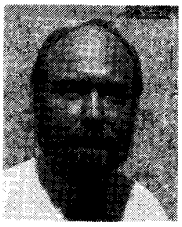
[22] ——, "Geometric-optical bidirectional reflectance modeling of a coniferous forest canopy," *IEEE Trans. Geosci. Remote Sensing*, vol. GE-24, no. 6, pp. 906-919, Nov. 1986.

[23] D. L. B. Jupp, J. Walker, and L. K. Penridge, "Interpretation of vegetation structure in Landsat MSS imagery: A case study in disturbed semi-arid eucalypt woodlands. Part 2. Model-based analysis," *J. Environ. Management*, vol. 23, pp. 35-57, 1986.

David L. B. Jupp received the B.Sc.(Hon) degree from Adelaide University and the Ph.D. degree from Flinders University (both in South Australia) in 1969 and 1973, respectively, majoring in applied mathematics.

He has worked in areas of geophysical inversion and remote sensing. He is currently a Principal Research Scientist with the CSIRO Division of Water Resources in Canberra, Australia. His work there involves remote-sensing applications in the water industry, commercial image processing systems development, and research into model-based interpretation of image data.

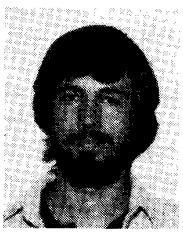
*



Alan H. Strahler (M'86) received the B.A. and Ph.D. degrees in geography from The Johns Hopkins University, Baltimore, MD, in 1964 and 1969, respectively.

He is currently Professor of Geography and Researcher in the Center for Remote Sensing, Boston University, Boston, MA. He has held prior academic positions at Hunter College of the City University of New York, at the University of California, Santa Barbara, and at the University of Virginia. Originally trained as a Biogeographer, he has been actively involved in remote sensing research since 1978. He has been a Principal Investigator on numerous NASA contracts and grants. His primary research interests are in spatial modeling and spatial statistics as they apply to remote sensing, and in geometric-optical modeling of remotely sensed scenes. He is particularly interested in remote sensing of forests and the inference of vegetation canopy parameters from digital images through invertible models.

*



Curtis E. Woodcock received the B.A., M.A., and Ph.D. degrees from the Department of Geography, University of California, Santa Barbara.

He is currently an Assistant Professor in the Department of Geography and Director of Geographic Applications in the Center for Remote Sensing at Boston University. His primary research interests in remote sensing include: mapping and inventory of forests and other natural environments, spatial analysis and pattern recognition, and image processing and geographic information systems. He is also interested in models of physical environments, particularly of vegetation canopies and geomorphic features in arid environments.

Dr. Woodcock is a member of the American Society of Photogrammetry and Remote Sensing.